Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

MAS

Modelling, Analysis and Simulation

Modelling, Analysis and Simulation

RKC time-stepping for advection-diffusion-reaction problems

J.G. Verwer, B.P. Sommeijer, W. Hundsdorfer

REPORT MAS-E0405 MARCH 23, 2004

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-3703

RKC time-stepping for advection-diffusion-reaction problems

ABSTRACT
The original explicit Runge-Kutta-Chebyshev (RKC) method is a stabilized second-order integration method for pure diffusion problems. Recently it has been extended in an implicit-explicit manner to also incorporate highly stiff reaction terms. This implicit-explicit RKC method thus treats diffusion terms explicitly and the highly stiff reaction terms implicitly. The current paper deals with the incorporation of advection terms for the explicit method, thus aiming at the implicit-explicit KRC integration of advection-diffusion-reaction equations in a manner that advection and diffusion terms are treated simultaneously and explicitly and the highly stiff reaction terms implicitly.

2000 Mathematics Subject Classification: 65M12, 65M20
1998 ACM Computing Classification System: G.1.1, G.1.7, G.1.8
Keywords and Phrases: Numerical integration, Runge-Kutta-Chebyshev methods, parabolic equations, hyperbolic equations, stiff ODEs, advection-diffusion-reaction problems.
Note: Work carried out within theme MAS1 and MAS3.

RKC time-stepping for advection-diffusion-reaction problems

J.G. Verwer, B.P. Sommeijer and W. Hundsdorfer
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

jan.verwer@cwi.nl, ben.sommeijer@cwi.nl, willem.hundsdorfer@cwi.nl

Abstract

The original explicit Runge-Kutta-Chebyshev (RKC) method is a stabilized second-order integration method for pure diffusion problems. Recently it has been extended in an implicit-explicit manner to also incorporate highly stiff reaction terms. This implicit-explicit RKC method thus treats diffusion terms explicitly and the highly stiff reaction terms implicitly. The current paper deals with the incorporation of advection terms for the explicit method, thus aiming at the implicit-explicit RKC integration of advection-diffusion-reaction equations in a manner that advection and diffusion terms are treated simultaneously and explicitly and the highly stiff reaction terms implicitly.

2000 Mathematics Subject Classification: Primary: 65M12, 65M20.

1998 ACM Computing Classification System: G.1.1, G.1.7 and G.1.8.

Keywords and Phrases: Numerical integration, Runge-Kutta-Chebyshev methods, parabolic equations, hyperbolic equations, stiff ODEs, advection-diffusion-reaction problems.
Note: Work carried out within themes MAS1 and MAS3.

1 Introduction

This paper is devoted to the time integration of stiff, nonlinear advection-diffusion-reaction equations. Adopting the method of lines approach we assume that the PDE system with its boundary conditions has been spatially discretized, and thus we focus our research on ODE systems

$$w'(t) = F(t, w(t)), \quad t > 0, \qquad w(0) = w_0, \tag{1.1}$$

representing semi-discrete advection-diffusion-reaction problems. In most practical applications the dimension of this ODE system is huge, especially for multi-space dimensional PDEs and/or PDE systems with many reacting species. The huge dimension and the simultaneous occurrence of advection, diffusion and reaction terms and stiffness can severely complicate the use of standard implicit integrators leaning on modified Newton and (preconditioned iterative) linear solvers. On the other hand, the stiffness induced by diffusion and reaction

terms rules out easy-to-use standard explicit solvers. This delineates our research question: how to realize easy-to-use, robust and efficient time stepping for this sort of semi-discrete PDEs.

Decoupling the three processes from one another generally simplifies matters. Most simple is to use operator (time) splitting by which advection, diffusion and reactions can be sequentially and independently solved with integrators tuned for the three different parts, see Ch. IV of [7]. A drawback is that operator splitting can give rise to large splitting errors for operators exhibiting slow and fast time scales that nearly balance. In particular, operator splitting is not exact for steady states which is a disadvantage for transient problems running into steady state. In this respect, decoupling through the implicit-explicit (IMEX) approach is more subtle and preserves transient balances.

In [18] we have proposed a Runge-Kutta-Chebyshev (RKC) method of the IMEX type treating modestly stiff diffusion terms explicitly and highly stiff reaction terms giving rise to real eigenvalues implicitly. The explicit method closely resembles the first RKC method due to van der Houwen & Sommeijer [6]. Here we examine our explicit method with the aim to also include advection terms. Our final goal is an efficient implicit-explicit RKC integration of advection-diffusion-reaction equations in a manner that advection and diffusion terms are treated simultaneously and explicitly and the highly stiff reaction terms implicitly.

2 The explicit RKC method

Historically the principal goal when constructing Runge-Kutta methods was to achieve the highest order possible with a given number of stages s. Stabilized methods like RKC are different in that only a few stages are used to achieve a usually low order whereas additional stages are exploited to increase the region of absolute stability, depending on the particular application. Originally the RKC method was intended for semi-discrete parabolic PDE problems. Correspondingly, the original method is stable on a strip containing a long segment of the negative real axis. The wider the strip, the greater the applicability of the method, but the most important characteristic of the formula is the length of the segment, the real stability boundary, which increases quadratically with s.

Let w_n denote the numerical approximation to the exact solution $w(t)$ of the semi-discrete system $w'(t) = F(t, w(t))$ at $t = t_n$ and let τ be the step size in the current step from t_n to t_{n+1}. The second-order explicit RKC formula has the form

$$\begin{aligned} W_0 &= w_n\,, \\ W_1 &= W_0 + \tilde{\mu}_1 \tau F_0\,, \\ W_j &= (1-\mu_j-\nu_j)W_0 + \mu_j W_{j-1} + \nu_j W_{j-2} + \tilde{\mu}_j \tau F_{j-1} + \tilde{\gamma}_j \tau F_0\,, \\ w_{n+1} &= W_s\,, \end{aligned} \tag{2.1}$$

where $j = 2, \ldots, s$. The W_k are internal vectors and F_k denotes $F(t_n + c_k\tau, W_k)$. All coefficients are available in analytical form for arbitrary $s \geq 2$:

$$\begin{aligned} &\tilde{\mu}_1 = b_1\omega_1 \quad \text{and for} \quad j = 2, \ldots, s\,, \\ &\mu_j = \frac{2b_j\omega_0}{b_{j-1}}\,,\ \ \nu_j = \frac{-b_j}{b_{j-2}}\,,\ \ \tilde{\mu}_j = \frac{2b_j\omega_1}{b_{j-1}}\,,\ \ \tilde{\gamma}_j = -a_{j-1}\tilde{\mu}_j\,, \end{aligned} \tag{2.2}$$

for which the a_j, b_j, c_j and ω_0, ω_1 are given below. Note the recursive form of W_j by which only 5 arrays of storage are needed for all $s \geq 2$.

When applied to the scalar stability test equation $w'(t) = \lambda w(t)$, we get at each stage a relation $W_j = P_j(z)W_0$ with $z = \tau\lambda$ and $P_j(z)$ a polynomial of degree j in z with $P_s(z)$ as stability function. Formula (2.1) has in fact been derived from a particular set of functions $P_j(z)\,(0 \leq j \leq s)$ satisfying three design criteria: (i) nearly optimal step-by-step stability of $P_s(z)$ for parabolic problems, (ii) internal stability, i.e., controlled round-off accumulation in a single step for s large, and (iii) second-order consistency of $P_j(c_j z)$ with respect to $e^{c_j z}$ for $j = 2, \ldots, s$. Criterion (iii) automatically implies second-order consistency of all $W_j\,(2 \leq j \leq s)$ at $t = t_n + c_j\tau$ for general problems $w'(t) = F(t, w(t))$. The first-stage formula is necessarily first-order consistent being forward Euler with step size $\tilde{\mu}_1\tau$.

The chosen functions P_j are based on the first kind Chebyshev polynomials $T_j(x)$ satisfying the three-term recursion

$$T_j(x) = 2x\,T_{j-1}(x) - T_{j-2}(x)\,, \qquad j = 2, 3, \ldots, s\,, \tag{2.3}$$

where $T_0(x) = 1$, $T_1(x) = x$. They are given by

$$P_j(z) = a_j + b_j T_j(\omega_0 + \omega_1 z)\,, \qquad a_j = 1 - b_j T_j(\omega_0)\,, \tag{2.4}$$

where [1)]

$$b_0 = b_2\,, \quad b_1 = 1/\omega_0\,, \quad b_j \;=\; T_j''(\omega_0)\,/\,(T_j'(\omega_0))^2\,, \qquad j = 2, \ldots, s\,, \tag{2.5}$$

with

$$\omega_0 = 1 + \epsilon/s^2\,, \qquad \omega_1 = T_s'(\omega_0)/T_s''(\omega_0)\,. \tag{2.6}$$

Here $\epsilon \geq 0$ is free and is called a damping parameter as $\varepsilon > 0$ gives values of the stability function $P_s(z)$ strictly less than one in the interior of the real stability interval. Later on we will exploit the freedom we have with ε to include advection terms.

Using $T_s'(1) = s^2, T_s''(1) = \frac{1}{3}s^2(s^2-1)$ and $T_s'''(1) = \frac{1}{15}s^2\,(s^2-1)\,(s^2-4)$, for ε small the stability boundary, denoted by $\beta(s)$, can be seen to satisfy [2)]

$$\beta(s) \;\approx\; \frac{2\omega_0\,T_s''(\omega_0)}{T_s'(\omega_0)} \;\approx\; \tfrac{2}{3}\,(s^2-1)\big(1 - \tfrac{2}{15}\epsilon\big)\,. \tag{2.7}$$

Taking $\epsilon = 2/13$, as in [7, 18], we get approximately $0.33 \leq P_s(z) \leq 0.95$ in most of the interior of the stability interval and a reduction in the $\beta(s)$ of about 2% to $\beta(s) \approx 0.65(s^2-1)$ compared to the undamped case ($\epsilon = 0$). Figure 2.1 illustrates the stability region $\mathcal{S} = \{z \in \mathbb{C} : |P_s(z)| \leq 1\}$ for $P_5(z)$ with and without damping. For larger values of s similar regions exist, except more stretched to the left along the negative real line.

Finally, by expanding $P_j(z)$ for $z \to 0$ it follows that the abscissa c_j are given by $c_j = b_j\omega_1 T_j'(\omega_1)$ and thus

$$c_0 = 0\,, \quad c_1 = c_2\,, \quad c_j = \frac{T_s'(\omega_0)}{T_s''(\omega_0)}\,\frac{T_j''(\omega_0)}{T_j'(\omega_0)}\,, \quad c_s \;=\; 1\,. \tag{2.8}$$

For ε small we then get $c_j \approx (j^2-1)/(s^2-1)$ for $2 \leq j \leq s-1$.

[1)] The choice for parameter b_1 differs from the choice $b_1 = b_2$ made in earlier RKC papers. The current choice was made in [18] to enable the IMEX form.

[2)] With ε small we actually mean $\varepsilon/s^2 \ll 1$. Likewise, if this does not hold we say that ε is large.

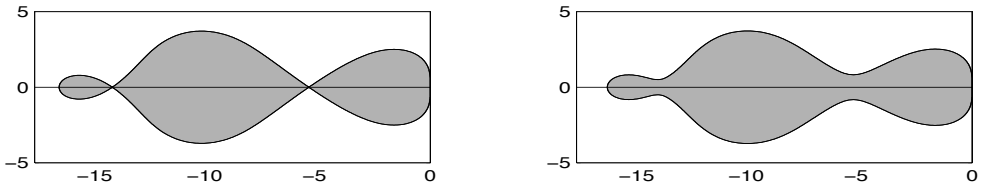

Figure 2.1: Stability regions for the second-order shifted Chebyshev polynomial P_5 with damping parameter ε small: left $\varepsilon = 0$, right $\varepsilon = 2/13$.

Remark 2.1 For $\epsilon = 0$ we have

$$P_s(z) = \frac{2}{3} + \frac{1}{3s^2} + \left(\frac{1}{3} - \frac{1}{3s^2}\right) T_s\left(1 + \frac{3z}{s^2-1}\right) \quad \text{with} \quad \beta(s) \approx \tfrac{2}{3}\,(s^2-1)\,.$$

This polynomial is due to Bakker, see [7], and generates about 80% of the optimal stability interval length for second-order polynomials, being $\beta(s) \approx 0.814 s^2$. Within most of the interior of the stability interval, $P_s(z)$ alternates between $\approx 1/3$ and 1. $\diamond$

Remark 2.2 The brief introduction in this section to RKC follows [18]. Related stabilized explicit methods are the ROCK [1, 2] and DUMKA methods [11, 12]. These have close to optimal real stability boundaries and can possess a higher order (up to order 4). The higher order makes them less amenable for the IMEX extension. Numerical comparisons between the 2-nd order RKC code from [16] (with still $b_1 = b_2$) and a 4-th order ROCK code can be found in [2, 7]. The IMEX version of this code has been used in [18]. More references are found in [7]. $\diamond$

Remark 2.3 Suppose system (1.1) can be split as

$$w'(t) = F_E(t, w(t)) + F_I(t, w(t))\,, \tag{2.9}$$

where F_I is too stiff to be treated efficiently by (2.1). The IMEX extension of (2.1) from [18] overcomes this for stiff terms F_I possessing a Jacobian matrix with a real spectrum. For problem (2.9) it reads, with j running from 2 to s,

$$\begin{aligned}
W_0 &= w_n\,,\\
W_1 &= W_0 + \tilde{\mu}_1 \tau F_{E,0} + \tilde{\mu}_1 \tau F_{I,1}\,,\\
W_j &= (1-\mu_j-\nu_j) W_0 + \mu_j W_{j-1} + \nu_j W_{j-2} + \tilde{\mu}_j \tau F_{E,j-1} + \tilde{\gamma}_j \tau F_{E,0} + \\
&\quad [\tilde{\gamma}_j - (1-\mu_j-\nu_j)\,\tilde{\mu}_1]\tau\, F_{I,0} - \nu_j \tilde{\mu}_1 \tau\, F_{I,j-2} + \tilde{\mu}_1 \tau\, F_{I,j}\,,\\
w_{n+1} &= W_s\,,
\end{aligned} \tag{2.10}$$

where $F_{E,j}$ denotes $F_E(t_n + c_j\tau, W_j)$, etc. As long as the Jacobian of F_I has a real spectrum, this method is unconditionally stable for the implicitly treated operator F_I and the stability is determined by the explicitly treated operator F_E, completely similar as we discussed above for method (2.1). The IMEX extension introduces a term to the $\mathcal{O}(\tau^3)$ local truncation error which is proportional to $\tau^2/(s^2-1)$. For s large this does no harm, otherwise accuracy reduction might be faced. See the analysis of [18] for details.

So in actual application the IMEX method is applied in the same manner as (2.1), except that we now encounter at each stage an implicit Euler type computation $W_j = W^* + \tilde{\mu}_1 \tau\, F_I(t_n + c_j\tau, W_j)$. If this implicit computation is cheap, as e.g. with stiff chemical reactions giving rise to small sized systems decoupled over space grids, the IMEX form is readily substantially more efficient than the fully explicit method applied to (2.9). In [18] only highly stiff diffusion-reaction problems were discussed. The adaptation of the explicit method towards advection-diffusion problems extends naturally to highly stiff advection-diffusion-reaction problems and the IMEX method. ◇

3 The link to advection-diffusion problems

The link to advection-diffusion problems is made through the damping parameter ε. Figure 2.1 illustrates that with a small $\varepsilon > 0$ the stability region $\mathcal{S}$ contains a narrow strip along the negative real line. By increasing ε the strip becomes wider, as illustrated in Figure 3.1 which shows $\mathcal{S}$ for P_5 for $\varepsilon = 5$ and $\varepsilon = \infty$. Obviously, by widening the strip eigenvalues with larger imaginary parts coming from advection terms can be put in. On the other hand, by increasing ε the strip also becomes shorter, meaning that less eigenvalues with large negative real part can be put in.

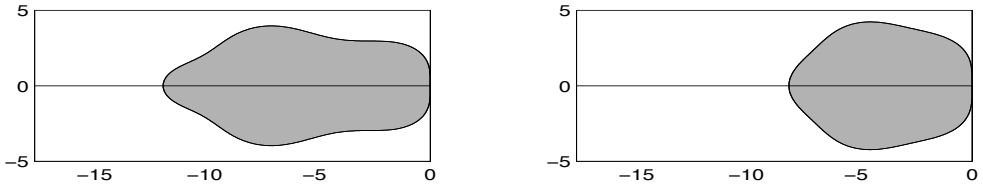

Figure 3.1: Stability regions for the second-order shifted Chebyshev polynomial P_5 with damping parameter ε large: left $\varepsilon = 5$, right $\varepsilon = \infty$.

3.1 The limit case $\varepsilon \to \infty$

There exists a surprising relation with a second-order scheme which in the numerical ODE field has been examined in connection with contractivity [9, 10] and in the numerical hyperbolic PDE field with respect to the TVD (Total Variation Diminishing) and SSP (Strong Stability Preserving) properties [4, 15, 17]. To show this we first examine the stability function (see (2.4))

$$P_s(z) \;=\; 1 + \frac{T_s''(\omega_0)}{(T_s'(\omega_0))^2}\Big(T_s(\omega_0 + \omega_1 z) - T_s(\omega_0)\Big) \tag{3.1}$$

for the limit case $\varepsilon \to \infty$ (only for $s \geq 3$ because for $s = 2$ there is no dependence on ε). For that purpose we use the representation

$$T_j(x) = \cosh(j\,\mathrm{acosh}(x)) = \frac{1}{2}\big(u^j + u^{-j}\big)\,, \qquad u = x + \sqrt{x^2 - 1}\,,$$

which holds due to $\mathrm{acosh}(x) = \ln(u)$, $x \geq 1$. Thus we have

$$T_j(x) \sim 2^{j-1}x^j\,, \quad T_j'(x) \sim j2^{j-1}x^{j-1}\,, \quad T_j''(x) \sim j(j-1)2^{j-1}x^{j-2}\,,$$

for $x \gg 1$. Inserting these asymptotic values for $j = s$ into (3.1) gives for $\varepsilon \to \infty$ the polynomial

$$K_s(z) \;=\; \frac{1}{s} + \frac{s-1}{s}\Big(1 + \frac{z}{s-1}\Big)^s . \tag{3.2}$$

The transition from P_s to K_s is quite surprising in the sense that K_s is precisely the second-order stability function obtained in [9] in a study of certain optimal linear monotonicity properties, and K_s is the stability function of the second-order explicit Runge-Kutta method

$$\begin{aligned} W_0 \;&=\; w_n\,, \\ W_j \;&=\; W_{j-1} \,+\, \tfrac{1}{s-1}\,\tau F\big(t_n + \tfrac{j-1}{s-1}\,\tau,\, W_{j-1}\big)\,, \quad j = 1, \ldots, s\,, \\ w_{n+1} \;&=\; \tfrac{1}{s}\,w_n \,+\, \tfrac{s-1}{s}\,W_s\,. \end{aligned} \tag{3.3}$$

which has been examined in [10] in a nonlinear contractivity study related to [9]. Furthermore, being based on a cyclic application of forward Euler, this method belongs to the class of TVD and SSP Runge-Kutta methods for hyperbolic problems [4, 15, 17].

So for linear problems the limit of our explicit RKC method (2.1) for $\varepsilon \to \infty$ is just method (3.3) as they share their stability function. They are also identical for $s = 2$ in the nonlinear case, being the explicit trapezoidal rule. For $s \geq 3$ the limit is then different though and given by

$$\begin{aligned} W_1 \;&=\; w_n + \tau^*\,F(t_n, w_n)\,, \\ W_2 \;&=\; \frac{1}{2}\,(w_n + W_1) + \frac{1}{2}\tau^*\,F(t_n + \tau^*, W_1)\,, \quad \text{and for } j = 3, \ldots, s\,, \\ W_j \;&=\; \frac{1}{j}\,W_0 - \frac{j-1}{j(j-2)}\,W_1 + \frac{(j-1)^2}{j(j-2)}\,\Big(W_{j-1} + \tau^*\,F(t_n + (j-2)\tau^*, W_{j-1})\Big)\,, \end{aligned} \tag{3.4}$$

where $\tau^* = \tau/(s-1)$ and $w_{n+1} = W_s$. By adjusting the parameter choice (2.5) to

$$b_j = 1/T_j(\omega_0) \quad (0 \leq j \leq s-1)\,, \qquad b_s = T_s''(\omega_0))/T_s'(\omega_0)\,, \tag{3.5}$$

the RKC formula (2.1) changes into

$$\begin{aligned} W_0 \;&=\; w_n\,, \qquad W_1 \;=\; w_n + \tilde{\mu}_1\tau\,F(t_n, w_n)\,, \\ W_j \;&=\; \nu_j\,W_{j-2} + \mu_j\,W_{j-1} + \tilde{\mu}_j\tau\,F(t_n + c_{j-1}\tau, W_{j-1})\,, \qquad j = 2, \ldots, s-1\,, \\ w_{n+1} \;&=\; (1 - \mu_s - \nu_s)\,w_n + \nu_s\,W_{s-2} + \mu_s\,W_{s-1} + \tilde{\mu}_s\tau\,F(t_n + c_{s-1}\tau, W_{s-1})\,, \end{aligned} \tag{3.6}$$

which does have (3.3) as limit for $\varepsilon \to \infty$. This formula maintains the stability function (3.1) and also the second-order consistency (although no longer at the internal stages). So for ε large (3.6) seems to be preferable for application to advection-diffusion problems.

Yet we do discard (3.6) since it has bad abscissa values c_j for ε small and we wish to use one and the same formula for all $\varepsilon > 0$ and all $s \geq 2$. For ε small we get, for $j = 1, \ldots, s-1$,

$$c_j = \frac{T_s'(\omega_0)T_j'(\omega_0)}{T_s''(\omega_0)T_j(\omega_0)} \approx \frac{3j^2}{s^2-1}\,,$$

and thus for ε small the values $t_n + c_j\tau$ can be far outside the time step interval $[t_n, t_{n+1}]$ for j close to s. This can be rather detrimental to accuracy and in fact this readily can be observed for problems with time-dependent Dirichlet boundary conditions.

An elementary calculation reveals that with the original parameter choice (2.5) we have abscissa satisfying $0 < c_1 = c_2 < c_3 < \cdots < c_j < \cdots < c_s = 1$ for all $\varepsilon \geq 0$. Since we consider this important we do prefer the original formula (2.1) for adjustment to advection-diffusion problems.

3.2 Fixing the damping parameter

The real stability boundary of (3.2) satisfies

$$\beta(s) \approx 2(s-1) \qquad \text{(exact for even } s\text{)}\,. \tag{3.7}$$

Hence the quadratic increase of $\beta(s)$ with s for ε small turns linear for $\varepsilon \to \infty$, showing that we rapidly lose stability for real negative eigenvalues if we take ε larger and larger. On the other hand, we then gain stability for eigenvalues with imaginary parts as already illustrated in Figure 3.1 for $s = 5$. For increasing s the stability region of (3.2) becomes circular with center point $1 - s$ and radius $s - 1$, so purely imaginary eigenvalues are excluded. This means that strongly advection dominated advection-diffusion problems solved with central differencing practically are out of reach. For such problems upwinding is to be preferred and this is even feasible in the pure advection case, see Section 3.3.

Because we do not wish to give up the quadratic increase with s of the real stability boundary (2.7) for diffusion dominated problems, the damping parameter ε cannot be chosen extremely large. In the remainder of the paper we fix ε to the value 10, unless noted otherwise. With $\varepsilon = 10$ the quadratic behaviour is maintained, viz.

$$\beta(s) \approx 0.34\,(s^2 - 1)\,,$$

for s large enough and ε still is large enough for including advection terms. Compared to the real stability boundary for $\varepsilon = 0$ we lose a factor two, approximately, which means a factor $\sqrt{2}$ for the number of function evaluations for strongly diffusion dominated problems. For implementation we will use the accurate approximation

$$\beta(s) = \begin{cases} 2\,, & s = 2\,, \\ (s^2-1)\left(0.340 + 0.189 \cdot (2/(s-1))^{1.3}\right), & s \geq 3\,. \end{cases} \tag{3.8}$$

3.3 The pure advection or diffusion case

The pure advection case is of interest in its own. Consider the test model $u_t + a u_x = 0$, assume periodicity in space with period one, and apply Fourier-von Neuman analysis for the third-order upwind-biased advection scheme. With CFL number $\nu = \tau|a|/h$ we then get the eigenvalues [7]

$$z = -\frac{4}{3}\nu \sin^4(\omega) - \frac{i}{3}\nu \sin(2\omega)\Big(4 - \cos(2\omega)\Big)\,, \qquad 0 \leq \omega \leq \pi\,.$$

For the RKC stability function (3.1) with $\varepsilon = 10$ and the stability function (3.2), Figure 3.2 shows plots of accurate estimates of the CFL limits $\nu(s)$ guaranteeing all $z \in \mathcal{S}$. For $s = 2$ they coincide and we have $\nu(2) \approx 0.87$. We see that for (3.2) the CFL limit $\nu(s)$ slowly increases with s. Clearly, the scaled value $\nu(s)/s$ is maximal for $s = 2$, implying that in the pure advection case this is the best choice with respect to efficiency (under the assumption

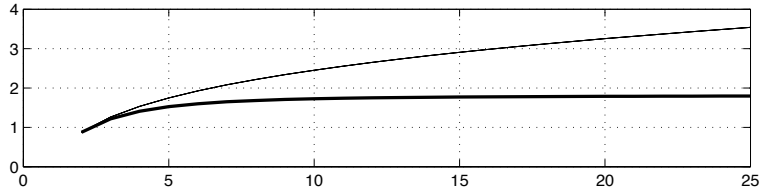

Figure 3.2: CFL limits for the stability functions (3.2) (solid line) and (3.1) with $\varepsilon = 10$ (fat solid line) for $s = 2, \ldots, 25$.

that the accuracy is practically independent of s which is true). The stability function (3.1) with $\varepsilon = 10$ gives somewhat smaller CFL limits, as expected. In this case $\nu(s)$ behaves constant for increasing s.

Observe that these CFL limits extend to the m-dimensional scalar test model

$$u_t + \sum_{k=1}^{m} a_k u_{x_k} = 0 \tag{3.9}$$

by summing up. For (3.1) with $\varepsilon = 10$ we then get

$$\sum_{k=1}^{m} \frac{\tau |a_k|}{h_k} \le \nu(s) = \begin{cases} \frac{4-s}{2}\, 0.87 + \frac{s-2}{2}\, 1.40\,, & 2 \le s \le 4\,, \\ \frac{9-s}{5}\, 1.40 + \frac{s-4}{5}\, 1.70\,, & 4 \le s \le 9\,, \\ 1.70\,, & s \ge 9\,, \end{cases} \tag{3.10}$$

where $\nu(s)$ now stands for an accurate lower bound of the CFL limit depicted in Figure 3.2. Needless to say that similar results can be obtained for the standard first-order and second-order upwind discretization. In the pure advection case the most efficient stable step size thus corresponds with the CFL limit for $s = 2$ and therefore it is advisable to restrict τ to be stable for $s = 2$.[3)]

In the pure diffusion case the situation is entirely different. With regard to stability we then put no bound on τ and s due the quadratic growth. In the pure diffusion case we thus can simply choose the minimal s satisfying the stability condition $\tau\sigma(F'(t,w)) = \beta(s)$, and this can be done for any given τ selected on the basis of accuracy considerations, e.g. by local error control as in the code from [16] (σ denotes here the spectral radius and $F'(t,w)$ is the Jacobian matrix).

To sum up, finding optimal critical step sizes for stability in the pure advection and the pure diffusion test model case is clear. However, for the mixed advection-diffusion test model case the situation is unclear and numerical stability analysis appears in general to be much more cumbersome.

4 Critical step sizes for advection-diffusion problems

Method of lines solvers for semi-discrete systems (1.1) are normally provided with variable step size control based on local error estimates. With such estimates one has a first tool at

[3)] For $s = 2$ formula (2.1) becomes $w_{n+1} = w_n + \frac{1}{2}\tau F(t_n, w_n) + \frac{1}{2}\tau F(t_{n+1}, w_n + \tau F(t_n, w_n))$, that is, the classical explicit trapezoidal rule which can be used profitably when combined with third-order upwind biased advection discretization, both limited and unlimited [7].

hand to timely prevent the onset of instabilities. The crucial question is can the step size control be trusted for this additional task. In the numerical ODE field research has been carried out in this direction under the names automatic stiffness detection and step-control stability, see Sect. IV.2 in [5] and Sect. 6.3 in [14].

For conditionally stable solvers it is natural to prescribe estimates of critical limits derived from stability analysis as maxima for the automatically chosen step sizes, provided they can be found with reasonable accuracy. For advection-diffusion problems an elegant approach is due to Wesseling, see [19] and Ch. V of [20]. For standard spatial discretizations of the m-dimensional scalar model

$$u_t + \sum_{k=1}^{m} a_k u_{x_k} = d \sum_{k=1}^{m} u_{x_k x_k} \,, \tag{4.1}$$

Wesseling gives step size conditions guaranteeing eigenvalues emerging from von Neumann stability analysis to lie inside geometric figures like squares, ellipses, half ellipses and ovals. For the integration method under consideration one then has to fit an appropriate figure inside the stability region $\mathcal{S}$ and to use the geometric step size condition to estimate the critical step size. For the RKC method ellipses and ovals seem suitable. Figure 4.1 shows regions $\mathcal{S}$ with an inscribed ellipse and oval for $s = 6$ and $\varepsilon = 0.1, 1, 10$. The numbers α, β represent the vertical half axis and the horizontal axis, respectively, the latter being equal to the real stability boundary (2.7). Observe the decrease of β and increase of α for increasing ε and also that α is smaller for the ovals. On the other hand, the ovals give a better fit near the origin which is important for advection dominated problems. In the remainder we therefore focus on the oval.[4)]

4.1 Oval step size conditions

Assume second-order central differencing for diffusion and the κ-scheme for advection with grid sizes h_k $(1 \le k \le m)$. For $\kappa = 1, -1$ and $1/3$ the κ-scheme yields, respectively, the second-order central, the second-order upwind and the third-order upwind-biased advection scheme. The parameterization of these standard schemes into the single κ-scheme is due to [13]. We consider the oval with center point $(-\beta/2, 0)$ and half-axes α and $\beta/2$, i.e.,

$$\Big(\frac{x}{\beta/2} + 1\Big)^2 + \Big(\frac{y}{\alpha}\Big)^4 = 1 \,.$$

As in Figure 4.1 we associate β with the real stability boundary $\beta(s)$. We then have one step size condition that emerges from the (artificial) diffusion terms, and one that emerges from the advection terms. The (artificial) diffusion step size condition is of course the familiar condition

$$\tau \le \frac{1}{2d \sum h_k^{-2} \, (2 + (1-\kappa) P_k)} \, \beta(s) \,, \tag{4.2}$$

where $P_k = |a_k| h_k / d$ is a mesh Péclet number. We emphasize that its violation will rapidly give instability for high-frequency error components. A condition like this thus is always imposed upon the RKC method and its use is not new. New is the advection step size

[4)] Results for the ellipse similar to those for the oval have been put in an appendix to this paper.

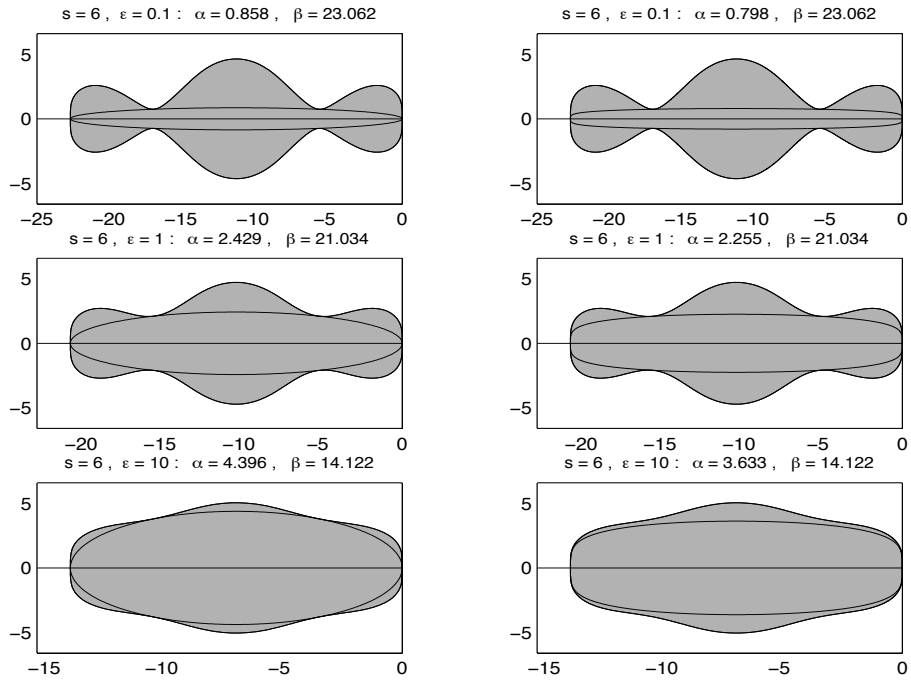

Figure 4.1: Stability regions $\mathcal{S}$ and inscribed ellipses (left) and ovals (right).

condition taken from [20] which reads

$$\tau \le q_1 \Big(\frac{4d\,\alpha^4(s)}{\beta(s)} \Big)^{1/3} \Big/ \sum \Big(\frac{a_k^4}{h_k^2} \Big)^{1/3}, \tag{4.3}$$

where the parameter q_1 depends on the choice of κ.[5)] For the popular κ-values, $q_1 \approx 0.635$ for $\kappa = 1/3$, $q_1 = 1$ for $\kappa = 1$, and $q_1 \approx 0.323$ for $\kappa = -1$. This condition generally will be conservative because the stability region $\mathcal{S}$ is not an oval and for the derivation of (4.3) the Cauchy-Schwarz inequality is used which is normally not sharp. But the proportionality with $d^{1/3}$ makes it interesting for small d, although it becomes meaningless for truly zero diffusion. With $q_1 \approx 0.635$ this advection condition is also applicable when the advection terms are discretized with the fourth-order central scheme [20].

Remark 4.1 In [20] one may choose between (4.3) and the CFL condition

$$\tau \le 2q_2 \, \frac{\alpha^2(s)}{\beta(s)} \Big/ \sum \frac{|a_k|}{h_k}, \tag{4.4}$$

[5)] This inequality is the corrected form of inequality (5.61) in [20], which contains an error. There the summation and taking the 1/3 power have been interchanged.

where $q_2 \approx 0.265$ for $\kappa = 1/3$, $q_2 = 0$ for $\kappa = 1$, and $q_2 = 0.317$ for $\kappa = -1$. Because the diffusion coefficient d is now absent, this CFL condition seems attractive for strongly advection dominated problems when using upwinding. However, it turns out that the coefficient $\alpha^2(s)/\beta(s)$ decreases with s and readily becomes too small for practical purposes. See the right plot of Figure 4.2 where for comparison also the plots for $\varepsilon = 0.1, 1, 3$ have been given. For this reason we discard this second oval condition. On the other hand, in Section 4.2 we will see that for $s = 2$ we have $\alpha^2/\beta \approx 1$, giving 0.53 as CFL limit for $\kappa = 1/3$ when using condition (4.4). The true critical CFL constant in this case equals approximately 0.87, see Figure 3.2. Hence for $s = 2$ and $\kappa = 1/3$ the oval estimate (4.4) is quite acceptable. $\diamond$

4.2 Optimal ovals

Since we prescribe the horizontal axis β by (2.7) we only have to compute the associated optimal half-axis $\alpha(s)$. Estimates for the optimal values for $\alpha(s)$ have been determined numerically. For $s = 2$ the optimal oval fit gives $\alpha(2) \approx \sqrt{2}$ and thus the ratio $\alpha^4(s)/\beta(s)$ needed in condition (4.3) equals 2, approximately. For $s \geq 3$ the semi-axis $\alpha(s)$ is an oscillating function of s with maxima for s even and minima for s odd. In the remainder we therefore restrict ourselves to even s. Further, for even s sufficiently large, $\alpha(s)$ becomes proportional to $\sqrt[4]{s^2-1}$ and thus the ratio $\alpha^4(s)/\beta(s)$ is then independent of s. The left plot of Figure 4.2 shows this ratio. For comparison also the plots for $\varepsilon = 0.1, 1, 3$ have been given.

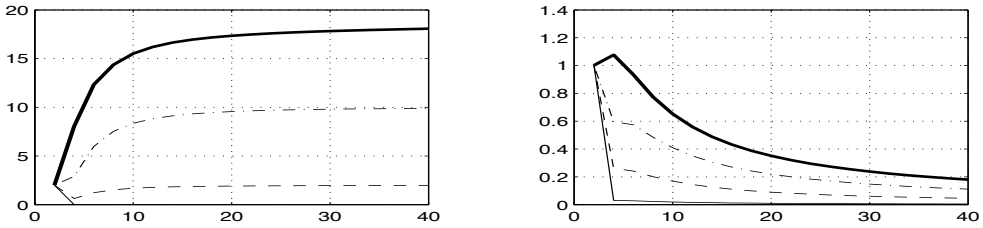

Figure 4.2: $\alpha^4(s)/\beta(s)$ (left) and $\alpha^2(s)/\beta(s)$ (right) for $\varepsilon = 0.1$ (solid), 1.0 (dashed), 3.0 (dash-dotted) and 10 (fat solid) for $s = 2, 4, \ldots, 40$.

For the actual implementation we will use the following lower bound for s even,

$$\frac{\alpha^4(s)}{\beta(s)} \approx \begin{cases} 2\,, & s = 2\,, \\ 4\,(6-s) + 6.15\,(s-4)\,, & s = 4, 6\,, \\ 6.15\,(10-s)/2 + 15.5\,(s-6)/4\,, & s = 8, 10\,, \\ 15.5\,, & s = 10, 12, \ldots\,. \end{cases}$$

Hence as maximum we take 15.5 which corresponds with $s = 10$. For larger values of s the slope in the curve becomes too small.

Remark 4.2 We will illustrate the oval conditions (4.2), (4.3) for the 1D model $u_t + au_x = du_{xx}$, using third-order upwind biased discretization. In terms of the CFL number $\tau|a|/h$ and the mesh Péclet number $P = h|a|/d$ we have

$$\frac{\tau\,|a|}{h} \leq \min\Big(q_1\Big(4\,\frac{\alpha^4(s)}{\beta(s)}\,\frac{1}{P}\Big)^{1/3},\ \frac{P}{2(2+(1-\kappa)P)}\,\beta(s)\Big)\,, \tag{4.5}$$

with $q_1 = 0.635, \kappa = 1/3$. Figure 4.3 plots the CFL limits based on the oval estimates for the Péclet numbers $P = 2, 10, 100$. For comparison also the associated true values are shown for the ovals (thus assuming that the stability regions are ovals) and for the stability regions themselves. The oval estimates are indeed conservative. On the other hand, the exact oval limits are rather good, especially for the advection dominated case. This is in line with the observation that the ovals do have a good fit with the stability regions near the origin. See Figure 4.1, case $\varepsilon = 10$. ◇

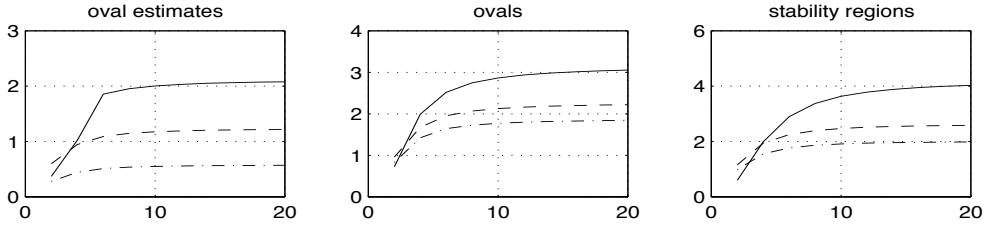

Figure 4.3: Third-order upwind-biased advection discretization: CFL limits for $s = 2, 4, \ldots 20$, for the Péclet numbers $P = 2$ (solid), 10 (dashed), 100 (dash-dotted).

4.3 Critical stepsize selection

Next suppose we are given a trial step size τ^* obtained from a local error estimation procedure. With this trial step size at hand we have the possibility to check the two stability inequalities (4.2), (4.3) and to adjust τ^* to a new step size τ to satisfy both inequalities. Simultaneously, the number of stages s with s even is to be adjusted such that the number of stages needed to satisfy the diffusion condition (4.2) is greater than or equal to the number of stages needed to satisfy the advection condition (4.3). This adjustment underlies the fact that the best strategy for advection is to minimize s and thus with respect to stability it makes no sense to spend more evaluations on advection than required by diffusion.[6)]

Let ψ_1 and ψ_2 contain the given problem parameters, i.e.

$$\psi_1 = \frac{1}{2d \sum h_k^{-2} \left(2 + (1-\kappa)P_k\right)}, \qquad \psi_2 = \frac{4d\, q_1^3}{\left(\sum (a_k^4/h_k^2)^{1/3}\right)^3}. \tag{4.6}$$

Then the following test is carried out:

1. If $\tau^* \le 2\psi_1$ we put $s = 2$, $\tau = \min\left(\tau^*, (2\psi_2)^{1/3}\right)$ and are done.
2. Put $\tau = \min(\tau^*, (15.5\,\psi_2)^{1/3})$. If $\tau \le 2\psi_1$ we put $s = 2$ and are done.
3. Determine $s_d \ge 4$ such that $\tau \le \beta(s_d)\psi_1$ to satisfy (4.2).
4. Determine $s_a \ge 4$ such that $\tau \le \left(\left(\alpha^4(s_a)/\beta(s_a)\right)\psi_2\right)^{1/3}$ to satisfy (4.3).

[6)] Standard for RKC is to impose the stability inequality (4.2) in the form of the more general condition $\tau\sigma(F'(t,w)) \le \beta(s)$ by only adjusting s, see [16]. Inequality (4.2) must be satisfied since its violation will amplify high-frequency error components with the possibility of overflow within a single integration step.

5. If $s_a \leq s_d$ we put $s = s_d$ and are done. Otherwise $\tau := 0.8\tau$ and we repeat steps 3, 4 and 5.

Remark 4.3 The above stability analysis is based on the test model (4.1). Nonlinear advection-diffusion systems such as $u_t + \nabla \cdot (\underline{a}u) = \nabla \cdot (D\,\nabla u)$ with $\underline{a} = \underline{a}(u)$ and $D = D(u)$ (positive diagonal), can be dealt with by applying the heuristic approach of 'freezing' as is customary in practice with von Neumann stability analysis. For the velocities a_k we then insert maximal values in ψ_1, ψ_2, and for the diffusion coefficient d a minimal value is required in ψ_2 and a maximal value in ψ_1.

One may also encounter pure advection coupled to mixed advection-diffusion or pure diffusion. Because for pure advection the oval approach is not applicable, one then should use the CFL condition (3.10) for the pure advection part of the problem. Thus, after the above oval test giving τ as new step size and s as the new number of stages, the step size adjustment

$$\tau := \min\Big(\tau, \nu(s)/\sum_{k=1}^{m} \frac{|a_k|}{h_k}\Big)\,.$$

is then to be carried out. If this adjustment is substantial one should iterate between the two tests to also adjust s. Recall that for pure advection alone the most efficient choice for the number of stages is $s = 2$. $\diamond$

5 Numerical illustrations

We will present numerical results obtained for two 3D test problems. In Section 5.1 a nonlinear Burgers type advection-diffusion problem is solved and in Section 5.2 we deal with an advection-diffusion-reaction problem with stiff reaction terms. For the advection-diffusion problem a modified version of the RKC solver from [16] has been used and for the second problem an IMEX extension thereof as discussed in Remark 2.3. Both of these are test solvers and not yet of the mature software level as the code from [16].

RKC is based on the second-order explicit scheme (2.1). The solver uses the damping parameter value $\varepsilon = 10$ instead of the standard value $\varepsilon = 2/13$ and expression (3.8) for the real stability boundary. RKC works as most other variable step size ODE solvers. A difference is that at each time step it minimizes the number of stages s so as to satisfy the stability condition $\tau\sigma \leq \beta(s)$, where σ is a spectral radius estimate for diffusion problems, such as the inverse of expression ψ_1 given in (4.6) coming from the first oval condition (4.2) which was used here. Variable step sizes are based on a local error per step criterion [14].

RKC has been used in two ways. (i) On the fly, that is, in the same way as for pure diffusion problems using only the first oval condition $\tau \leq \psi_1\beta(s)$, and not being protected to instability caused by advection terms, and (ii) also protected to instability caused by advection terms through the additional oval condition (4.3). In case (ii) the step size strategy of Section 4.3 has been used.

Once ψ_1, ψ_2 and Tol have been prescribed the required step size and stage tests go automatically. Standard the Euclidean norm is used for the local error test. Of course, for a given value of Tol the maximum norm would more timely signal the onset of instabilities.

5.1 A 3D Burgers type problem

We consider the three-space dimensional problem

$$u_t + \frac{1}{2}\,(u^2)_x + (\frac{3}{2}\,u - \frac{1}{2}\,u^2)_y + (\frac{3}{2}\,u - \frac{1}{2}\,u^2)_z = d\,\Delta u \tag{5.1}$$

in the unit cube on the time interval $[0, 1]$. This problem is a nice test model for nonlinear advection-diffusion and is derived from the 3D Burgers equation [3]. It admits the exact wave front solution

$$u(x, y, z, t) = 1 - \frac{1}{2}\left(1 + e^{(-x+y+z-3t/4)/(4d)}\right)^{-1} \tag{5.2}$$

which moves skew in the cube. This exact solution has been used to prescribe Dirichlet boundary conditions.

We have conservatively discretized on a uniform space grid with third-order upwind-biased for advection and second-order central for diffusion, using the grid sizes $h = 1/50$, $1/100$, $1/200$ and the diffusion coefficients $d = 10^{-2}, 10^{-3}, 10^{-4}$ (nine cases were tested). The true solution varies between 0.5 and 1.0. So freezing coefficients yields as maximal velocities for the stability test model (4.1) the values $a_1 = a_2 = a_3 = 1$, which can now be used to estimate the maximal scaled step sizes τ/s imposed by the oval conditions (4.2) and (4.3). Figure 5.1 shows these scaled maxima for six test cases for the even numbers of stages $s = 2, 4, \ldots, 20$. The point where the o and $*$ markers intersect determines the optimal oval-based values for τ and s. The step size strategy of Section 4.3 should determine these oval-based values automatically (in close approximation), and has done this right in the runs discussed below. Note that the advection oval condition predicts $s = 2$ already for $d = 10^{-3}$ on all three grids (advection dominated). Of course the same happens for $d = 10^{-4}$ (not shown here).

5.1.1 Test results

The above described solver RKC has been applied with values of Tol ranging from 10^{-1} to 10^{-4} (smaller values are less appropriate since the order is only two) over the time interval $[0, 1]$. On the fly, that is without the oval condition, it consistently (all nine test cases) ran into instability for $Tol = 10^{-1}$, so here the local error control failed to timely detect the onset of instabilities. For $Tol = 10^{-2}$ it consistently produced stable and accurate results, but occasionally with somewhat more step rejections than normal, see Table 5.1. For $Tol = 10^{-3}$ and 10^{-4} all on the fly integrations were successful too and now with very few step rejections. Summarizing, on the fly the solver works normal and efficient and for the current problem there appears to be little need to safeguard its control by means of the oval condition.

Of course, imposing this condition is safer. Indeed, then the instabilities for $Tol = 10^{-1}$ do not occur and the solver automatically selects the oval-based step sizes and numbers of stages predicted in Figure 5.1, but with a higher expense in integration steps and function calls. To illustrate this, we have collected integration data and L_2-errors at time $t = 1$ in Table 5.1 and Table 5.2 for, respectively, $Tol = 10^{-2}$ and 10^{-3} (being typical values for a second order solver like RKC). The results are given for the $100 \times 100 \times 100$ grid (results on the two other grids are similar). The entry Steps represents the accepted plus the rejected integration steps. The errors are the full PDE errors and are strongly dominated by their

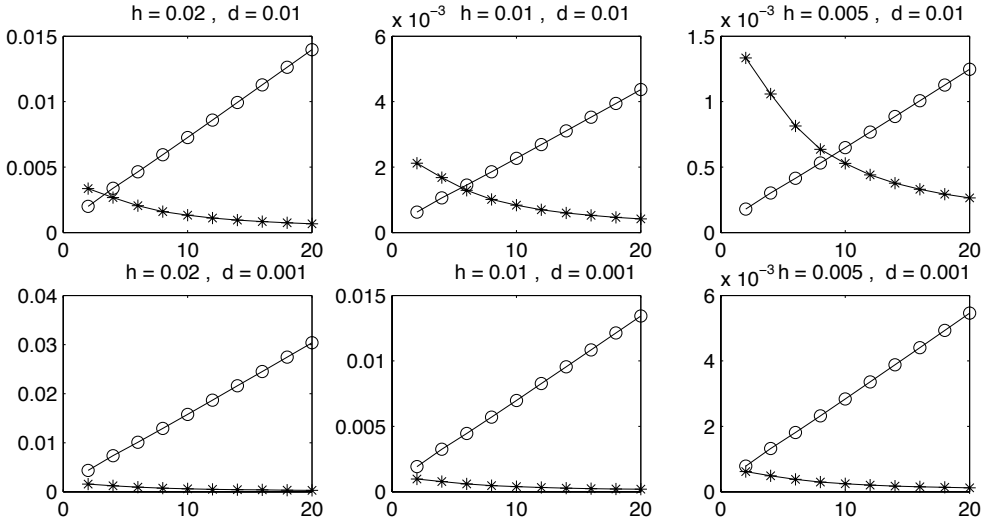

Figure 5.1: Problem (5.1). Scaled maximal step sizes τ/s according to the oval conditions (4.2) (-o-marker) and (4.3) (-$*$- marker) plotted as a function of $s = 2, 4, \ldots, 20$.

spatial parts. Hence they are not to be understood as ODE time integration errors (often these are smaller than the spatial errors).

$Tol = 10^{-2}$, $h = 10^{-2}$	Steps (rej)	F-evals	s_{max}	L_2-error
$d = 10^{-2}$: On the fly	30 (4)	413	25	$0.24\,10^{-3}$
Oval condition	150 (0)	899	6	$0.46\,10^{-4}$
$d = 10^{-3}$: On the fly	133 (14)	508	9	$0.37\,10^{-2}$
Oval condition	327 (0)	656	2	$0.38\,10^{-2}$
$d = 10^{-4}$: On the fly	153 (1)	476	5	$0.93\,10^{-2}$
Oval condition	593 (0)	1188	2	$0.59\,10^{-2}$

Table 5.1: Problem (5.1). Results for $Tol = 10^{-2}$.

$Tol = 10^{-3}$, $h = 10^{-2}$	Steps (rej)	F-evals	s_{max}	L_2-error
$d = 10^{-2}$: On the fly	29 (2)	390	19	$0.27\,10^{-3}$
Oval condition	150 (0)	899	6	$0.46\,10^{-4}$
$d = 10^{-3}$: On the fly	120 (1)	478	5	$0.36\,10^{-2}$
Oval condition	327 (0)	656	2	$0.38\,10^{-2}$
$d = 10^{-4}$: On the fly	166 (1)	498	3	$0.57\,10^{-2}$
Oval condition	593 (0)	1188	2	$0.59\,10^{-2}$

Table 5.2: Problem (5.1). Results for $Tol = 10^{-3}$.

5.2 A 3D advection-diffusion-reaction problem

We will next illustrate the IMEX version of RKC mentioned in Remark 2.3. For this purpose we consider a two-component, 3D advection-diffusion-reaction problem of the form

$$u_t + a_1 u_x + a_2 u_y + a_3 u_z = d\Delta u + f(u)\,. \tag{5.3}$$

The velocities a_k are given scalars, d is a given constant diffusion coefficient, $u = [u_1, u_2]^T$, and $f(u)$ is a stiff, nonlinear reaction term with components

$$f_1(u) = -k_2 u_1 u_2 + k_1 u_2^2\,, \qquad f_2(u) = -k_1 u_2^2 + k_2 u_1 u_2\,,$$

where k_1, k_2 denote given positive constants. As space domain we take the unit cube, as initial time $t = 0$ and as end time for output $t = 1$. Due to the stiff reaction term positivity is essential for this problem since negative solution values (wiggles) can easily result in instability or breakdown of the modified Newton iteration in stiff reaction computations. In spite of their simplicity, the chosen reaction terms reveal this.

To sketch the solution behaviour we first consider the case $d = 0$ as in [3]. Solutions then can be interpreted as solutions of the reaction part along characteristics of the advection operator. The reaction part has the general solution

$$u_1(t) = \frac{s_0}{k_1 + k_2}\,\frac{k_1(1-\alpha) + (k_1+k_2)\alpha e^{-s_0 k_2 t}}{1 - \alpha + \alpha e^{-s_0 k_2 t}}\,, \quad u_2(t) = s_0 - u_1(t)\,, \tag{5.4}$$

where $\alpha = \big((k_1+k_2)u_1(0) - s_0 k_1\big)/s_0 k_2$ and $s_0 = u_1(t) + u_2(t)$ which is constant in time. We choose $u_1(0) = 0, u_2(0) = s_0$ and introduce stiffness by putting $k_1 = k_2 = k \gg 1$. Component $u_1(t)$ then very rapidly increases from 0 to $s_0/2$ and likewise $u_2(t)$ rapidly decreases from s_0 to $s_0/2$. After the transient, the stiff eigenvalue of the reaction Jacobian is close to $-ks_0$ (the other eigenvalue is equal to zero). In the remainder we put $k = 10^6$.

For the a_k and pure advection solution we choose, following [3],

$$\begin{aligned} &a_1 = \pi\sqrt{2}\,(y + z - 1)\,, \quad a_2 = -\pi\sqrt{2}\,(x - 1/2)\,, \quad a_3 = a_2\,,\\ &u_{adv}(x,y,z,t) = \exp\big(-80\big[\big(x - r(t)\big)^2 + \big(y - s(t)\big)^2 + \big(z - s(t)\big)^2\big]\big)\,, \end{aligned} \tag{5.5}$$

where $r(t) = \big(2 + \sin(2\pi t)\big)/4$ and $s(t) = \big(4 + \sqrt{2}\cos(2\pi t)\big)/8$. These a_k define a rotation with period one along the characteristics (e.g. ellipses in the plane $y = z$) and the profile can

be visualized as a 3D plume with highest values equal to one along the curves defined by $r(t)$ and $s(t)$. As solution $u(x, y, z, t)$ for the advection-reaction problem we thus have (5.4) with s_0 replaced by $u_{adv}(x, y, z, t)$, getting zero as initial function for u_1 and the initial profile from (5.5) as initial function for u_2. Both u_1 and u_2 will rapidly approach $u_{adv}(x, y, z, t)/2$ and after the transient the solution changes in time only by advective transport.

The numerical tests have been carried out with $d > 0$. Then no exact solution is known, but for d very small the behaviour will be alike. As initial values we have used the initial values from the advection-reaction solution and these initial values were also used to prescribe Dirichlet boundary values for $t \in [0, 1]$.

For space discretization we have used the same approach as for problem (5.1), except that here the third-order upwind-biased advection scheme has been provided with flux limiting to prevent unwanted negative solutions. We have used the max-min limiter from [8], see also Section III.1.1 in [7]. The spatial discretization thus results in an ODE system of type (2.9) where F_E contains the advection-diffusion terms and F_I the reactions. Due to the limiting the function F_E is strongly nonlinear, but since F_E is treated explicitly this renders no problem.

The IMEX-RKC formula (2.10) has been implemented in the variable step size solver briefly discussed in the beginning of Section 5, see also [18]. This solver has been applied in precisely the same way as the explicit solver was applied to problem (5.1) (cases (i) and (ii) described in the beginning of Section 5). The main difference is that here also implicit reaction computations are to be performed. For the current problem these implicit computations take about 1/3 of the total CPU time. Recall that these implicit computations are decoupled over the space grid and hence can be dealt with by a standard modified Newton process as is customary in stiff ODE computations.

5.2.1 Test results

Table 5.3 contains results for the diffusion coefficient $d = 10^{-6}$, thus numerically mimicking the advection-reaction case with $d = 0$. The data in the table is similar to the data given in Tables 5.1, 5.2, with as L_2-error the full error with respect to the exact advection-reaction solution. Because we are numerically dealing with advection-reaction the CFL stability step size adjustment of Remark 4.3 has been used rather than the oval condition. Obeying the available CFL condition is advocated to avoid significant negative values which can ruin the integration process. As advocated in Remark 4.3, the number of stages s has been taken equal to 2 so that according to (2.10) the integration formula applied here is just the following IMEX form of the explicit trapezoidal rule,[7)]

$$\begin{aligned} W_1 &= w_n + \tau F_E(t_n, w_n) + \tau F_I(t_{n+1}, W_1)\,, \\ w_{n+1} &= \tfrac{1}{2}(w_n + W_1) + \tfrac{1}{2}\tau F_E(t_{n+1}, W_1) - \tfrac{1}{2}\tau F_I(t_n, w_n) + \tau F_I(t_{n+1}, w_{n+1})\,. \end{aligned}$$

All runs were successful. The results in the table are for the local error tolerance $Tol = 10^{-3}$. The tolerance values $10^{-1}, 10^{-2}$ gave nearly the same results. Apparently, most of the time the CFL condition overrules the local error control. Note that the CFL condition restricts the step size τ to $\tau \leq 0.87h/(2\pi\sqrt{2})$. With a constant step size this would have resulted in 511 and 1022 steps, respectively. The solver requires a few more steps due to the initial transient phase. The L_2-error is mainly spatial.

[7)] This formula is not the most efficient IMEX extension of the trapezoidal rule. For the current illustration it suffices however.

$Tol = 10^{-3}$	Steps (rej)	F-evals	s_{max}	L_2-error
$h = 2.0\,10^{-2}$	539 (1)	1080	2	$0.26\,10^{-2}$
$h = 10^{-2}$	1048 (1)	2098	2	$0.47\,10^{-3}$

Table 5.3: Problem (5.3) with $d = 10^{-6}$. CFL condition added to step size control.

Table 5.4 contains results for the diffusion coefficient values $d = 10^{-1}, 10^{-2}, 10^{-3}$ obtained on the $100 \times 100 \times 100$ grid for $Tol = 10^{-3}$ (L_2-errors cannot be given now and the comments also apply to $Tol = 10^{-1}, 10^{-2}$ and/or the $50 \times 50 \times 50$ grid). We have applied the RKC-IMEX solver with the oval condition imposed on the local error control, as in Section 5.1. All runs were completed succesfully. Note that with $d = 10^{-3}$ we come close to the advection-reaction data in Table 5.3, telling us that the oval condition does a fairly good job here.

$Tol = 10^{-3}$	Steps (rej)	F-evals	s_{max}
$d = 10^{-1}$	269 (1)	3442	14
$d = 10^{-2}$	678 (1)	2658	4
$d = 10^{-3}$	1151 (1)	2304	2

Table 5.4: Problem (5.3). Oval condition added to step size control.

Next we have repeated the tests of Table 5.4 on the fly without any protection for instabilities due to advection and thus only relying on step size control based solely on the local error estimate and with s sufficiently large to satisfy (4.2). These integrations failed due to significant negative values, resulting in Newton divergence and even overflow within a single integration step. They confirm that with the combination of advection terms and stiff reactions great care must be exercised with step control stability based solely on a local error estimate. In this regard it is clear that the oval condition offers more robustness.

After a simple tentative negativity test was added to the step size control, all these on the fly runs became successful too, see Table 5.5. The negativity test was carried out at all stages and allows step rejection. Negativity was concluded when a stage component came below -10^{-10}. This results in step rejection and halving the current step size. Also the maximal growth factor for τ was lowered from 10 to 2 for safety. In Table 5.5 only the accepted steps with the associated numbers of function evaluations have been listed in view of the preliminary character of the negativity test. In terms of CPU time these tentative on the fly runs compare with the oval runs of Table 5.4, which certainly is an asset of the oval condition. Finally we note that resetting negative values to zero was not used since this would interfere with the mass balance.

In a sequel to this paper we plan to upgrade the current test version of the IMEX solver to a software tool providing the same level of robustness and reliability as the explicit solver from [16] designed for pure diffusion problems. Tables 5.3, 5.4 and 5.5 clearly indicate that such an upgrade will result in an efficient and reliable advection-diffusion-reaction solver.

$Tol = 10^{-3}$	Steps	F-evals	s_{max}	$\text{CPU}_{\text{fly}}/\text{CPU}_{\text{oval}}$
$d = 10^{-1}$	274	2672	44	1.13
$d = 10^{-2}$	230	1236	9	0.76
$d = 10^{-3}$	372	1270	4	0.86

Table 5.5: Problem (5.3). On the fly integration with negativity test. Only accepted steps have been counted in view of the preliminary character of the negativity test.

6 Concluding remarks

In this paper we have demonstrated how the explicit s-stage RKC method originally proposed for diffusion problems [6] can be adjusted for advection-diffusion problems by simply resetting the damping parameter ϵ. The method then can efficiently integrate with high order upwind CFL limits near 1 and no limitation on s to cope with (moderately stiff) diffusion terms. For $s = 2$ the method is just the explicit trapezoidal rule which we advocate for pure advection problems. Hence the scope of the explicit method ranges from diffusion dominated to advection dominated. Together with the IMEX extension from [18] to include severely stiff reaction terms, we thus have got a new method suitable for integrating a wide class of advection-diffusion-reaction problems. An attractive feature is that the advection-diffusion computations are explicit and that the reaction computations are decoupled over the space grid.

Finding critical time step sizes for advection-diffusion problems for use in actual solvers is a stability problem on its own. For the RKC method we have demonstrated the geometric approach of [19, 20], using the oval condition. This condition clearly enhances robustness, but the resulting step size values can be too restrictive. A possible alternative is improved step control stability [5, 14], tuned for advection-diffusion-reaction problems. With the obtained experience in mind, in the near future we plan to upgrade the current test version of the IMEX solver to a validated and mature piece of software, similar as the existing explicit RKC code [16].

Acknowledgements The authors acknowledge support from the national program BSIK: knowledge and research capacity. J.G. Verwer and B.P. Sommeijer for the BRICKS subproject AFM2-3 (Numerical algorithms in bioinformatics) and W. Hundsdorfer for the BRICKS project MSV-1 (Scientific computing).

References

[1] A. Abdulle, A.A. Medovikov (2001), *Second order Chebyshev methods based on orthogonal polynomials*. Numer. Math. 90, pp. 1–18.

[2] A. Abdulle (2002), *Fourth order Chebyshev methods with recurrence relation*. SIAM J. Sci. Comput. 23, pp. 2042–2055.

[3] J.G. Blom, J.G. Verwer (1994), *VLUGR3: a vectorizable adaptive grid solver for PDEs in 3D. I. Algorithmic aspects and applications*. Appl. Numer. Math. 16, no. 1-2, pp. 129–156.

[4] S. Gottlieb, C.-W. Shu, E. Tadmor (2001), *Strong stability preserving high-order time discretization methods*. SIAM Review 42, pp. 89-112.

[5] E. Hairer, G. Wanner (1996), *Solving Ordinary Differential Equations II – Stiff and Differential-Algebraic Problems*. Second edition, Springer Series in Computational Mathematics, Vol. 14, Springer, Berlin.

[6] P.J. van der Houwen, B.P. Sommeijer (1980), *On the internal stability of explicit, m-stage Runge-Kutta methods for large m-values*. Z. Angew. Math. Mech. 60, pp. 479–485.

[7] W. Hundsdorfer, J.G. Verwer (2003), *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Series in Computational Mathematics, Vol. 33, Springer, Berlin.

[8] B. Koren (1993), *A robust upwind discretization for advection, diffusion and source terms.* In: *Numerical Methods for Advection-Diffusion Problems*. Eds. C.B. Vreugdenhil, B. Koren, Notes on Numerical Fluid Mechanics, Vol. 45, Vieweg, Braunschweig, pp. 117-138.

[9] J.F.B.M. Kraaijevanger (1986), *Absolute monotonicity of polynomials occurring in the numerical solution of initial value problems*. Numer. Math. 48, pp. 303–322.

[10] J.F.B.M. Kraaijevanger (1991), *Contractivity of Runge-Kutta methods*. BIT 31, pp. 482–528.

[11] V.I. Lebedev (1994), *How to solve stiff systems of differential equations by explicit methods*. In: *Numerical Methods and Applications*. Ed. G.I. Marchuk, CRC Press, pp. 45–80.

[12] V.I. Lebedev (2000), *Explicit difference schemes for solving stiff problems with a complex or separable spectrum*. Comput. Math. and Math. Phys. 40, pp. 1801–1812.

[13] B. van Leer (1985), *Upwind-difference methods for aerodynamic problems governed by the Euler equations*. In: *Large Scale Computations in Fluid Mechanics*. Eds. B.E. Engquist, S. Osher, R.C.J. Somerville, AMS Series, American Mathematical Society, pp. 327–336.

[14] L.F. Shampine (1994), *Numerical Solution of Ordinary Differential Equations*. Chapman & Hall, New York.

[15] C.-W. Shu, S. Osher (1988), *Efficient implementation of essentially non-oscillatory shock-capturing schemes*. J. Comput. Phys. 77, pp. 439–471.

[16] B.P. Sommeijer, L.F. Shampine, J.G. Verwer (1997), *RKC: An explicit solver for parabolic PDEs*. J. Comput. Appl. Math. 88, pp. 315–326.

[17] R.J. Spiteri, S.J. Ruuth (2002), *A new class of optimal high-order strong-stability-preserving time-stepping schemes*. SIAM J. Numer. Anal. 40, pp. 469-491.

[18] J.G. Verwer, B.P. Sommeijer (2003), *An implicit-explicit Runge-Kutta-Chebyshev scheme for diffusion-reaction equations*. Report MAS-R0305, preprint CWI, to appear in SIAM J. Sci. Comput.

[19] P. Wesseling (1996), *Von Neumann stability conditions for the convection-diffusion equation*. IMA J. of Num. Anal. 16, pp. 583-598.

[20] P. Wesseling (2001), *Principles of Computational Fluid Dynamics*. Springer Series in Computational Mathematics, Vol. 29, Springer, Berlin.

Appendix on the ellipse conditions

In this appendix we list some results for the ellipse with center point $(-\beta/2, 0)$ and half-axes α and $\beta/2$, i.e.,

$$\Big(\frac{x}{\beta/2}+1\Big)^2+\Big(\frac{y}{\alpha}\Big)^2=1\,.$$

These results have been obtained in the same experimental way as the results for the oval. Similar as for the oval, for the ellipse we have one step size condition that emerges from the (artificial) diffusion terms and one that emerges from the advection terms. The (artificial) diffusion step size condition is identical for the ellipse and the oval, being the familiar condition (4.2) with $\beta(s)$ approximated by (3.8) (we again assume $\varepsilon = 10$). For the ellipse the advection condition reads [20]

$$\tau \le \frac{4d}{(2-\kappa)^2\sum a_k^2}\,\frac{\alpha^2(s)}{\beta(s)}\,. \tag{6.1}$$

If $\kappa = 1$ (second-order central differencing), then together with (4.2) this condition is sufficient and necessary for the eigenvalues emerging from von Neumann analysis to lie inside the ellipse with semi-axes $\alpha(s)$ and $\beta(s)/2$. Recall that for second-order central differencing the eigenvalue curve itself is an ellipse. For $\kappa \neq 1$ this condition merely is sufficient and generally is conservative because the stability region $\mathcal{S}$ is not ellipse shaped and Cauchy-Schwarz type inequalities are involved in its derivation. In particular, for $d \to 0$ the step size has to go to zero linearly with d. Apparently, condition (6.1) is then less suitable than the oval condition (4.3) in view of the proportionality with $d^{1/3}$ in (4.3). With $q_1 \approx 0.635$ this advection condition is also applicable when the advection terms are discretized with the fourth-order central scheme.

Optimal ellipses

We need to estimate optimal values for the ratio $\alpha^2(s)/\beta(s)$. For $s = 2$ we have $\alpha(2) = \sqrt{3}$, $\beta(2) = 2$ so that $\alpha^2(2)/\beta(2) = 3/2$. We found experimentally that for s sufficiently large $\alpha(s)$ is proportional to $\sqrt{s^2-1}$. Since $\beta(s)$ behaves as s^2-1, the ratio $\alpha^2(s)/\beta(s)$ is then independent of s. Figure 6.1 illustrates this for $\varepsilon = 0.1, 1, 3, 10$. It nicely reveals that $s = 2$ is a separate case and that for increasing values of ε the asymptotic behaviour requires increasing values of s to hold. For actual implementation one can use the lower bounds

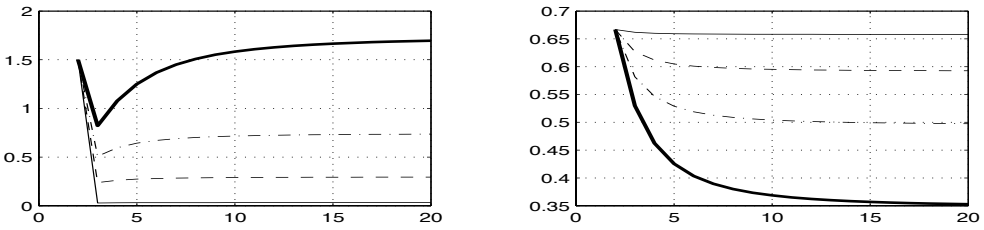

Figure 6.1: $\alpha^2(s)/\beta(s)$ (left) and $\beta(s)/(s^2-1)$ (right) for $\varepsilon = 0.1$ (solid), 1.0 (dashed), 3.0 (dash-dotted) and 10 (fat solid) as a function of $s = 2, \ldots, 20$.

$$\frac{\alpha^2(s)}{\beta(s)} = \begin{cases} \frac{3}{2}\,, & s = 2\,, \\ \frac{8-s}{5}\,\frac{82}{100} + \frac{s-3}{5}\,\frac{3}{2}\,, & 3 \le s \le 8\,, \\ \frac{3}{2}\,, & s \ge 8\,. \end{cases} \tag{6.2}$$

We have chosen 3/2 as maximum which corresponds with $s = 8$. Allowing more stages makes no sense since the slope of the curve rapidly goes to zero for $s > 8$. The ellipse conditions can now be implemented in a critical step size selection strategy in precisely the same way as we have done for the oval.

Remark 6.1 Similar to Remark 4.2 it is instructive to illustrate the ellipse condition based on (4.2) and (6.1) for the 1D model $u_t + au_x = du_{xx}$ in terms of the CFL number $\tau|a|/h$ and the mesh Péclet number $P = h|a|/d$. An elementary calculation gives

$$\frac{\tau|a|}{h} \le \min\Big(\frac{4}{(2-\kappa)^2}\,\frac{1}{P}\,\frac{\alpha^2(s)}{\beta(s)}\,,\; \frac{P}{2(2+(1-\kappa)P)}\,\beta(s)\Big)\,. \tag{6.3}$$

For second-order central advection discretization ($\kappa = 1$), Figure 6.2 plots these CFL limits for $P = 2, 10, 100$ and for comparison the associated true values as well. The figure confirms that the ellipse condition is only practically feasible for diffusion dominated problems with small to moderate Péclet numbers. For large values the condition is too restrictive. Also note that for $\kappa = 1/3$ (third-order upwind biased) the limits become even less favourable.$\diamond$

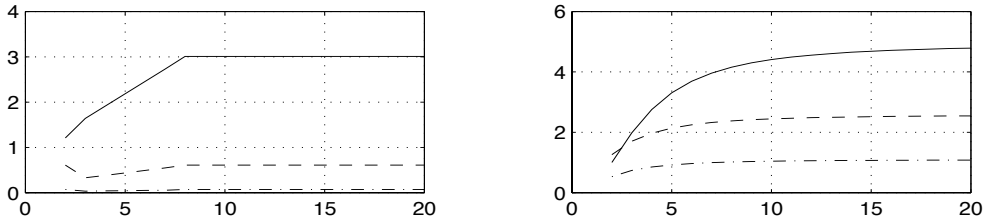

Figure 6.2: Second-order central advection discretization: CLF numbers based on the ellipse condition (left) and associated true values (right). Plotted as function of $s = 2, \ldots, 20$ for the Péclet numbers $P = 2$ (solid), 10 (dashed), 100 (dash-dotted).